



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Molecular Epidemiology and Evolution of Human Respiratory Syncytial Virus and Human Metapneumovirus

Citation for published version:

Gaunt, ER, Jansen, RR, Poovorawan, Y, Templeton, KE, Toms, GL & Simmonds, P 2011, 'Molecular Epidemiology and Evolution of Human Respiratory Syncytial Virus and Human Metapneumovirus', *PLoS ONE*, vol. 6, no. 3, e17427. <https://doi.org/10.1371/journal.pone.0017427>

Digital Object Identifier (DOI):

[10.1371/journal.pone.0017427](https://doi.org/10.1371/journal.pone.0017427)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

PLoS ONE

Publisher Rights Statement:

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Molecular Epidemiology and Evolution of Human Respiratory Syncytial Virus and Human Metapneumovirus

Eleanor R. Gaunt^{1*}, Rogier R. Jansen², Yong Poovorawan³, Kate E. Templeton⁵, Geoffrey L. Toms⁴, Peter Simmonds¹

1 Centre for Infectious Diseases, University of Edinburgh, Edinburgh, United Kingdom, **2** Department Medical Microbiology, Academic Medical Centre, Amsterdam, Netherlands, **3** Center of Excellence in Clinical Virology, Chulalongkorn University, Bangkok, Thailand, **4** Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, United Kingdom, **5** Specialist Virology Centre, Royal Infirmary of Edinburgh, Edinburgh, United Kingdom

Abstract

Human respiratory syncytial virus (HRSV) and human metapneumovirus (HMPV) are ubiquitous respiratory pathogens of the *Pneumovirinae* subfamily of the *Paramyxoviridae*. Two major surface antigens are expressed by both viruses; the highly conserved fusion (F) protein, and the extremely diverse attachment (G) glycoprotein. Both viruses comprise two genetic groups, A and B. Circulation frequencies of the two genetic groups fluctuate for both viruses, giving rise to frequently observed switching of the predominantly circulating group. Nucleotide sequence data for the F and G gene regions of HRSV and HMPV variants from the UK, the Netherlands, Bangkok and data available from Genbank were used to identify clades of both viruses. Several contemporary circulating clades of HRSV and HMPV were identified by phylogenetic reconstructions. The molecular epidemiology and evolutionary dynamics of clades were modelled in parallel. Times of origin were determined and positively selected sites were identified. Sustained circulation of contemporary clades of both viruses for decades and their global dissemination demonstrated that switching of the predominant genetic group did not arise through the emergence of novel lineages each respiratory season, but through the fluctuating circulation frequencies of pre-existing lineages which undergo proliferative and eclipse phases. An abundance of sites were identified as positively selected within the G protein but not the F protein of both viruses. For HRSV, these were discordant with previously identified residues under selection, suggesting the virus can evade immune responses by generating diversity at multiple sites within linear epitopes. For both viruses, different sites were identified as positively selected between genetic groups.

Citation: Gaunt ER, Jansen RR, Poovorawan Y, Templeton KE, Toms GL, et al. (2011) Molecular Epidemiology and Evolution of Human Respiratory Syncytial Virus and Human Metapneumovirus. PLoS ONE 6(3): e17427. doi:10.1371/journal.pone.0017427

Editor: Ron Fouchier, Erasmus Medical Center, Netherlands

Received: November 5, 2010; **Accepted:** February 2, 2011; **Published:** March 1, 2011

Copyright: © 2011 Gaunt et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by a studentship awarded to E.G. by the BBSRC, the Thailand Research Fund, the Commission on Higher Education, King Chulalongkorn Memorial Hospital, the Center of Excellence in Clinical Virology and CU Centenary Academic Development Project, Chulalongkorn University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: e.gaunt@sms.ed.ac.uk

Introduction

Human respiratory syncytial virus (HRSV) and human metapneumovirus (HMPV) are globally ubiquitous respiratory pathogens of the *Pneumovirinae* subfamily of the *Paramyxoviridae*. Both viruses comprise two genetic groups, A and B, distinguishable genetically and serologically [1–3] which co-circulate with fluctuating frequencies. The two HRSV genetic groups are referred to as subgroups; these comprise genotypes distinguished on the basis of antibody cross reactivity [4] or phylogeny [5]. Each of the two HMPV genetic groups are referred to somewhat paradoxically as genotypes, and each genotype comprises two sub-genotypes (A1, A2, B1 and B2). HMPV genotypes are distinguishable serologically and sub-genotypes are discerned phylogenetically [3].

Fluctuating circulation frequencies of HRSV subtypes and HMPV genotypes give rise to the observation of switching of the predominantly circulating subtype (HRSV) or genotype (HMPV) between respiratory seasons [6–13]. HMPV was discovered in 2001 and so longitudinal epidemiologic studies are infrequent,

though for HRSV a theme of cyclicity whereby subtype A predominates for a number of seasons then subtype B predominates (usually for a shorter duration) are reported [14–19]. HRSV-A is considered the major subtype in terms of both frequency [20] and associated morbidity [7]. Similarly, HMPV-A strains are generally detected at a higher frequency than HMPV-B strains [12,13,21] and clinical differences are reported between HMPV genotypes [21,22]. Repeat HRSV infections occur throughout life with decreasing morbidity, and increasingly evidence suggests the same is also true for HMPV [3,23–27].

The HRSV and HMPV virions both express two highly immunogenic surface proteins against which adaptive immune responses are directed. The fusion (F) protein mediates fusion of viral and cell membranes and is highly conserved. Anti-HRSV antibody directed against F protein is cross-reactive for strains of both subtypes [1,2,28,29], and studies on HMPV using human sera and animal models have indicated similar antibody reactivity patterns [30–33]. It is conceivable that the conformational changes arising on activation of the fusion protein [34] serve to expose the conserved functional (and immunogenic) regions (analogous to the

gp41 fusion protein of HIV-1 [35]) which are otherwise, in the native state, sheltered from immunologic recognition.

The attachment (G) glycoprotein of pneumoviruses conversely portrays several immune evasive traits. Specificity of antibody raised against the G protein extends to, and possibly beyond, the genotype level (HRSV) or sub-genotype level (HMPV) [4,29,30,36–38]. In both viruses the G protein is extensively glycosylated with both *N*- and *O*-linked sugars and a high proportion of proline residues [39,40] thought to reduce ordered secondary structure of the protein [41].

The highly variable G protein of HRSV comprises 298 [42] or 318 [43] residues. Two mucin-like hypervariable regions (HVRs) at the C terminus under significant positive selection form hydrophilic stalk like protrusions from the surface of the virion separated by an exposed [44] but conserved and non-glycosylated region comprising residues 151–190 [45]. A heparin binding domain (HBD) identified between residues 184–198 (subgroup A) or 183–197 (subgroup B) binds heparin-like glycosaminoglycans (GAG) on the host cell surface [46]. A fourteen residue region incorporating four universally encoded cysteine residues at positions 173, 176, 182 and 186 believed to form disulphide bridges is common to human and bovine RSVs [47], and the G protein has been shown to bind host cell receptor CX3CR1 via the CX3C chemokine domain accommodating cysteine residues at positions 182 and 186 [48]. Interestingly, subtype specific seroconversion directed against this antigenic region is detectable in only 40% of individuals [49].

The HMPV G protein, which comprises 217 to 236 residues [50,51] has not been resolved in such detail. Nevertheless, great variability in the C terminal ectodomain is seen; conversely to HRSV, no conserved cysteine pairs or chemokine domains are detectable. HMPV does not encode a conserved methionine or alternative initiation codon in or adjacent to the transmembrane region which would permit production of a secreted form of the G protein, unlike HRSV [52].

It is hypothesized that switching of the predominant circulating subtype of HRSV is brought about by short-lived subtype-specific herd immunity in a population generated over one or two seasons, which favours dissemination of the alternate subtype in a subsequent season [43,53,54]. This suggestion has been borrowed to explain HMPV genotype switching also [12].

Evolutionary modelling of several respiratory viruses has been undertaken. A well conducted analysis which identified positively selected sites and the evolutionary characteristics of HRSV was bipartite, corresponding with the two subtypes A and B [10,11]. Sequence data spanning 47 and 45 years (for subtypes A and B respectively) were used to identify positively selected sites within the attachment (G) protein, to distinguish residues which had a significant likelihood of being *O*-glycosylated (a mechanism used to shelter residues from immunologic recognition [55]), and to determine the time since the most recent common ancestor (t_{MRCA}) of the HRSV species – which was estimated to have existed 350 years ago.

The evolutionary dynamics of the closely related HMPV have also been explored [56–58]. The t_{MRCA} of the HMPV sub-genotypes were estimated at 12–28 years, and the two genotypes were found to have mean t_{MRCA} s of 26–51 years [56–58]. The t_{MRCA} across the species in one study was 119–133 years [56], whereas another study using sequence data collected over a greater number of years determined a more recent species level t_{MRCA} of 97 years [58], though both analyses were conducted using similar sequence data and the same software package (BEAST [59]).

Many human infecting viruses, both in the field of respiratory medicine and more widely, undergo a turnover and replacement

of predominant lineages with emergent strains. For example, clades of echoviruses 9, 11 and 30 are frequently replaced by novel recombinant forms with striking periodicity [60], dengue virus serotypes are comprised of clades undergoing replacement [61], and measles virus clade replacement is described, the latter of which is likely due to selective advantages brought about by the vaccine era [62]. HRSV and HMPV have not previously been discussed in such terms, and so it was decided to investigate whether similar evolutionary mechanisms are evident for HRSV and HMPV.

We report identification and evolutionary modelling of three geographically dispersed contemporary clades of HRSV and five of HMPV, which have circulated for decades. Switching of the predominantly circulating genotype (HMPV)/subtype (HRSV) therefore cannot be attributed to the emergence of novel virus lineages. Identification of numerous sites under positive selection in the G proteins of both viruses were frequently discordant with those identified in previous studies, and little overlap in positively selected sites was observed between HRSV subtypes or HMPV genotypes. The interpretations of these findings are discussed.

Results

Circulating clades of HRSV and phylogenetic reconstructions

To investigate the geographical and temporal distribution of individual clades of each virus, phylogenetic analyses were performed on 243 HRSV and 310 HMPV F gene sequences with isolation dates spanning 44 and 26 years respectively (Fig. 1). Three clades of HRSV with sufficient sequences available were identified (labelled 1–3; Fig. 1A), of which two were subtype A, and one was subtype B.

HRSV subtype B sequences largely grouped into the one clade, and all but three subtype B sequences falling outside this clade were collected before 2002. HRSV subtype A sequences mostly fell into one of two clades, with all but one of the sequences not belonging to one of the identified clades having a collection date prior to 2002. There was little evidence of geographical clustering of HRSV sequences, with strains from Newcastle, the Netherlands and Bangkok phylogenetically interspersed among the Edinburgh strains in all three clades (Fig. 1A). HRSV clade 1 was comprised entirely of sequences generated during this study, whereas clades 2 and 3 incorporated sequences downloaded from Genbank of Asian origin.

Phylogenetic analyses comparing the same 58 HRSV isolates sequenced in the F and G gene regions (Fig. 2) reveals congruence between the two datasets with sequence clusters supported by a 70% bootstrap threshold consistent over the two genome regions. (Greater bootstrap support was evident in the G gene, reflective of the greater diversity seen in this region.) Closer inspection of the HRSV-A F and G gene phylogenies revealed two monophyletic lineages (Fig. 2). Directional evolution of the two HRSV-A lineages is evident visually.

Identification of positively selected sites in the HRSV genes encoding surface proteins

To further understand the evolutionary pressures acting on HRSV and HMPV, analyses to detect codons under positive selection in the F and G genes of both viruses were undertaken. No positively selected codons were identified in the F gene of HRSV clades, findings that contrasted markedly with the 32 positively selected sites detected within the HRSV-A G gene and 5 in the HRSV-B G gene (Table 3). The codon encoding residue 258 was the only residue identified as positively selected for both subtypes.

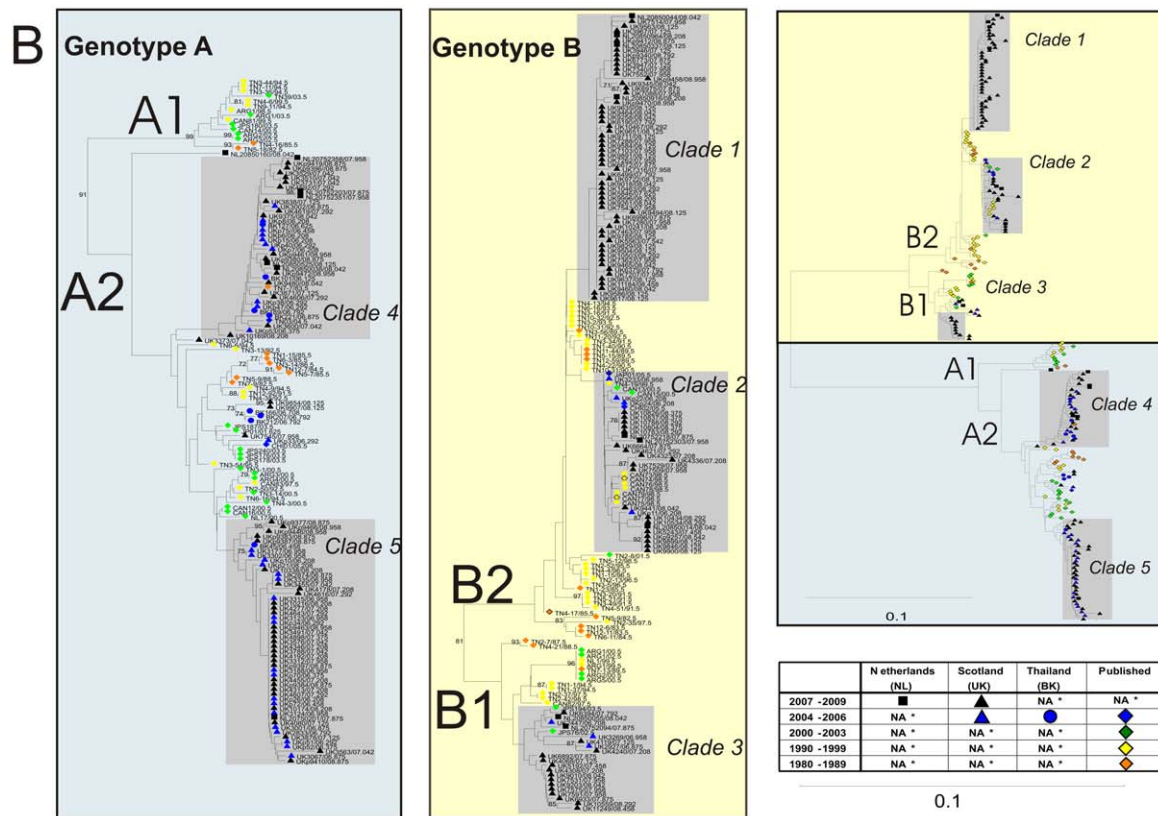
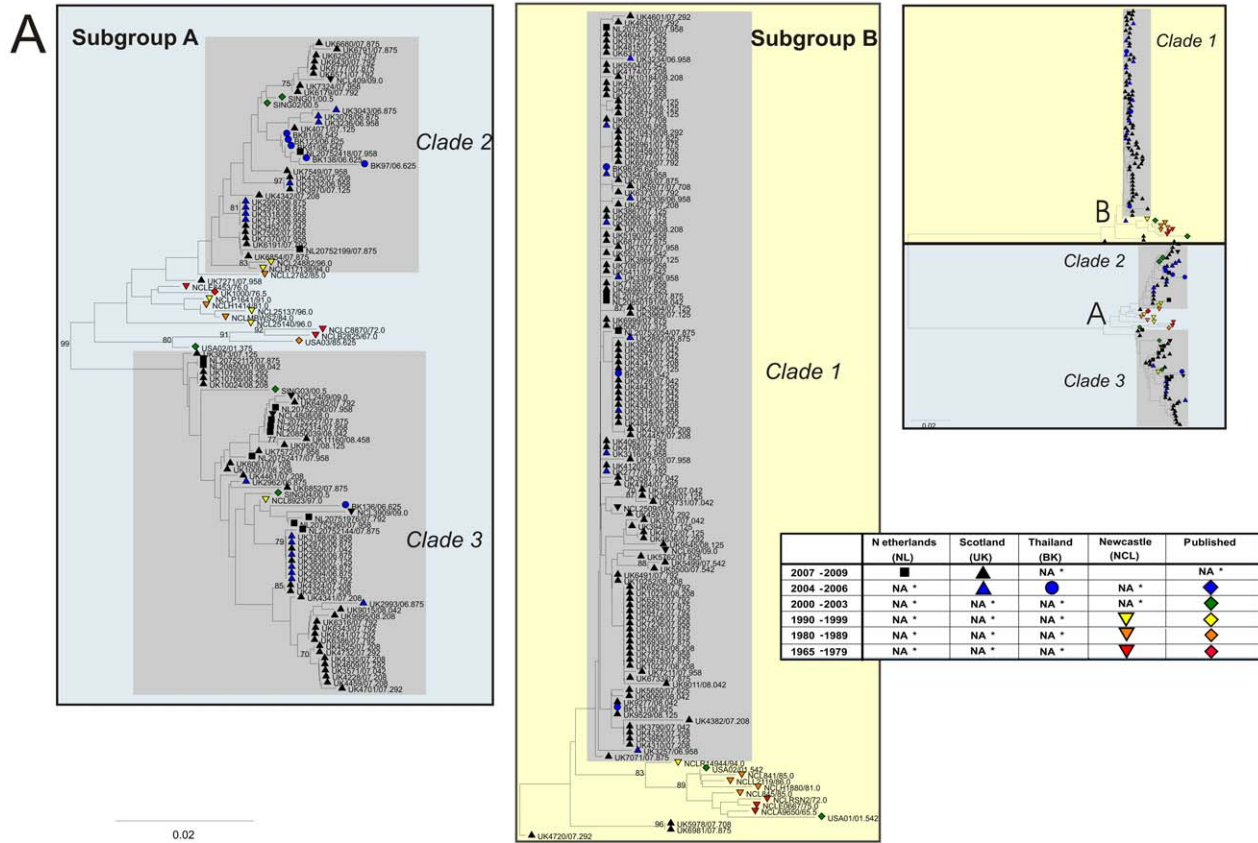


Figure 1. Phylogenetic analysis of HRSV (A) and HMPV (B) partial F gene sequences. Phylogenetic reconstruction was by neighbour joining of MCL-corrected pair-wise distances. Clades identified as described in the methods are indicated by grey shaded boxes. Sequence symbols are colour coded by year of isolation. Symbol shape denotes geographic origins sequences. Bootstrap values >70% are indicated. (A) Phylogenies rooted with bovine RSV (not shown). Subgroups A and B are indicated by the blue and yellow boxes respectively. (B) rooted with avian metapneumovirus species C (not shown). Genotypes A and B are indicated by the blue and yellow boxes respectively. NA, None analysed.
doi:10.1371/journal.pone.0017427.g001

Circulating clades of HMPV

HMPV phylogenetic reconstructions revealed 5 main monophyletic groups corresponding with five major clades, one within sub-genotype B1, and two each within sub-genotypes A2 and B2 (labelled 1–5; Fig. 1B). The availability of geographically diverse HMPV-F sequences in Genbank allowed identification of strains from at least two, and up to four continents within clades 2, 3, 4 and 5. Clade 1 was the exception, comprised entirely of strains from Edinburgh and the Netherlands. Older HMPV strains clustered to the internal nodes of the tree. HMPV G protein sequence data available in Genbank spanning 11 years between 1997 and 2008 was analysed phylogenetically (Fig. 3). This also revealed geographically disparate strains interspersed phylogenetically, and that several identified lineages circulated concurrently, confirming the observations drawn from the F gene phylogenetic analysis, and comparable with HRSV. Together with the strong evidence of directional evolution, this is indicative of epidemiologic and evolutionary traits shared by the two members of the *Pneumovirinae*.

HMPV strains from Bangkok were exclusively of genotype A, though previously determined strains of Japanese origin grouped in genotype B (clades 2 and 3). One unusual HMPV genotype A sequence from the Netherlands (NL20850160/08.042) did not group into either sub-genotype A1 or A2. Despite sampling from three globally distributed referral centres collected over two years, we were unable to identify any sequences belonging to the HMPV A1 sub-genotype.

Positive selection in the HMPV surface proteins

Analysis of the F gene of HMPV-B sub-genotypes yielded only two positively selected codons in the B2 group at residues 391 ($p = 0.870$) and 400 ($p = 0.664$). Within the G gene of HMPV-A, fourteen codons were positively selected, six were positively selected within HMPV-B1 and 17 sites were identified as under positive selection within HMPV-B2 (Table 3). For HMPV-A, HMPV-B1 and HMPV-B2, sites identified as under positive selection were usually different between groups, except residue 110 which was identified in all three groups.

HRSV-F and HMPV-F clade turnover

Evolutionary analyses of three HRSV-F and five HMPV-F clades was undertaken to determine the minimum length of time monophyletic lineages of both viruses have been circulating (Table 4). Strict molecular clock models were always used under the assumption that the evolutionary rate within a clade did not vary. Analysis of the HRSV-B clade (clade 1, Fig. 1A) was unable to yield an ESS>200, likely due to the low diversity encompassing most F gene sequences within the subtype and so was excluded from analyses. Indeed, in this clade, visual evidence of directional evolution was not evident from phylogenetic analysis. The *tMRCA* of HRSV-A clades was 17 and 14 years and of HMPV clades was between 11 and 27 years. Genetic diversity across HRSV and HMPV clades was around 2–3% (Table 4). Congruent *tMRCA*, diversity and substitution rates of clades of both viruses (Table 4) support the existence of evolutionary mechanisms common to both virus species.

Discussion

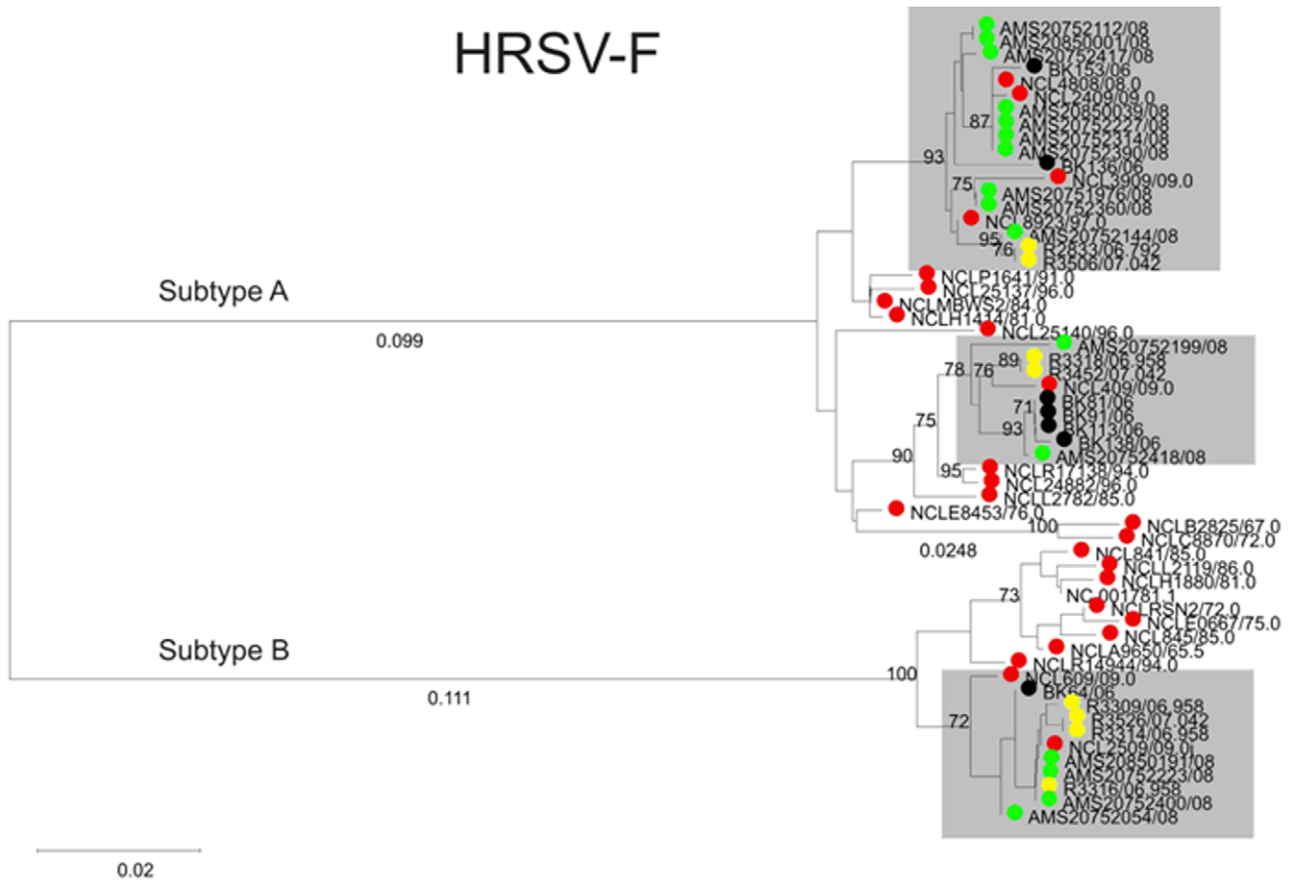
HRSV and HMPV sequence data generated from isolates collected over 44 years and 26 years respectively were analysed using a variety of techniques, including phylogenetic reconstruction, evolutionary modelling and identification of positively selected sites to gain insight into the epidemiology and evolution of these closely related viruses. Identification of contemporary HRSV and HMPV clades was undertaken for evolutionary modelling, to further understanding of the circulation trends of predominant virus lineages. Evolutionary analyses of the *Pneumovirinae* have not previously been approached in this way.

The evolutionary rates of F gene sequences for HMPV clades in the range of $1.0\text{--}1.7 \times 10^{-3}$ substitutions/site/year (Table 4) are slightly higher, albeit within the 95% highest posterior density (HPD) intervals, than the rates calculated in previous studies sampling across the species of 0.9×10^{-3} [56] and 0.712×10^{-3} substitutions/site/year [57]. It has previously been noted that external branches of the HMPV-F phylogenetic reconstruction have higher dN/dS ratios than internal branches [56]. Taken together, these observations suggest that substitutions are more frequently selected for in the contemporary virus population than previously. This might be explained by an increasing virus population size – random sampling of a larger virus population increases the likelihood of detection of nucleotide changes, and increasing population size increases the probability that residue changes will be selected for.

Previous calculation of the evolutionary rate of HRSV has been undertaken by analysing G gene sequence data, with rates of 1.83×10^{-3} and 1.95×10^{-3} substitutions/site/year determined for subtypes A and B respectively [10,11]. These rates are slightly higher than those calculated here for the F protein ($1.3\text{--}1.5 \times 10^{-3}$ substitutions/site/year). A higher evolutionary rate in the G protein than the F protein has similarly been observed for HMPV [56]. The G protein is under evolutionary pressure due to the host population adaptive immune response to this immunogenic region [45,63–66]. Nucleotide substitutions, most commonly those which are non-synonymous, are frequently selected for, and as nucleotide changes become fixed in the population they are more likely to be captured by the evolutionary analyses, which might explain why the model yields higher rates in this region. Conversely, the extremely low dN/dS ratio and lack of positively selected sites in the F protein [67] provides evidence that changes in the F protein are deleterious, probably due to functional constraints.

For both HMPV and HRSV, we have identified an abundance of sites under positive selection within the G gene, but few within the F gene. In HRSV-A, we identified 32 sites under positive selection. Previous analysis to detect positively selected sites within HRSV-A [10] identified twelve positively selected sites using sequence data spanning a similar time frame, with 48 sequences analysed compared with 139 here; 9 residues were identified by both analyses (111, 117, 215, 226, 262, 274, 276, 290, 297). The previous study used the same program and threshold for significance, and similar models and sampling frame in terms of the number of representative strains and date range analysed. An explanation of this incongruence might be in the differences between the predominantly circulating lineages of the

HRSV-F



HRSV-G

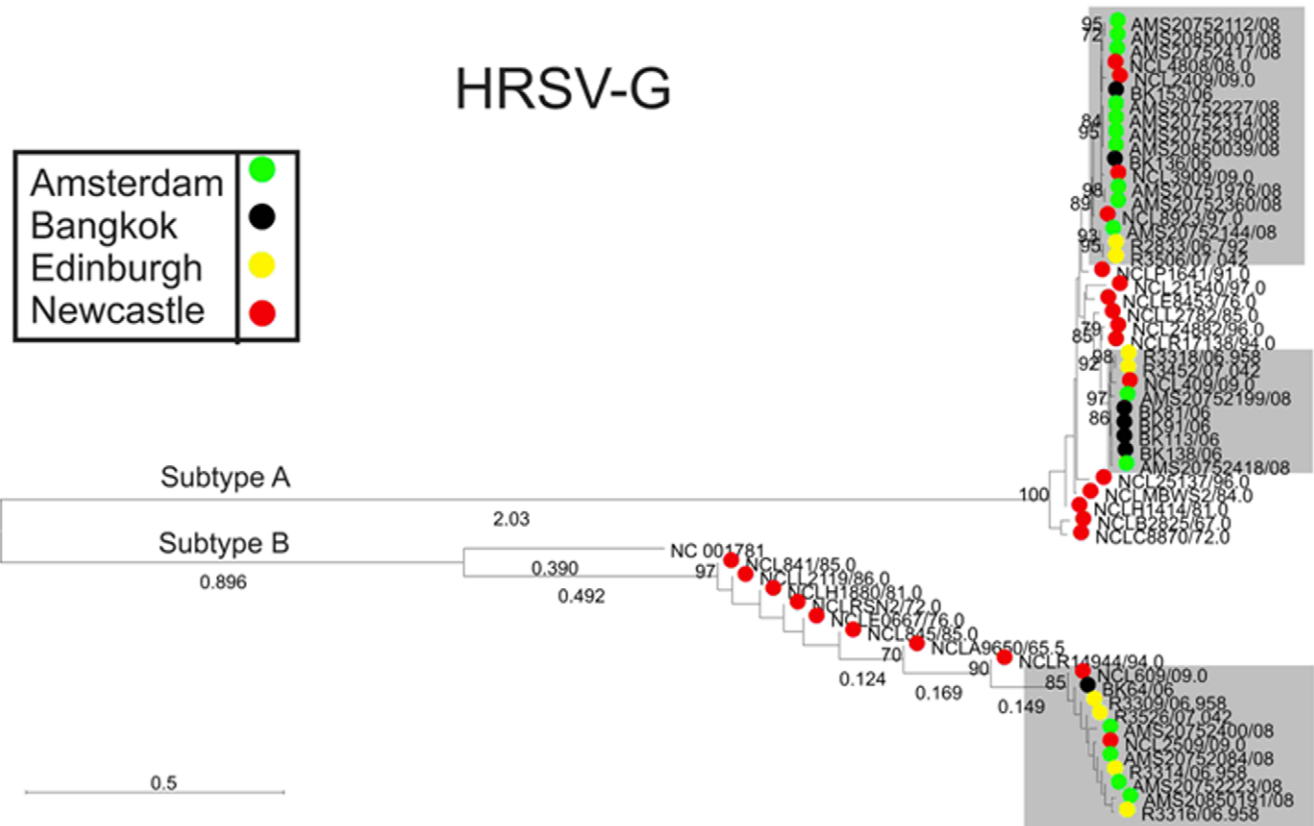


Figure 2. Phylogenetic analysis of 58 HRSV sequences in the F and G genes. Phylogenetic reconstruction was by neighbour joining of MCL-corrected pair-wise distances. F gene sequences rooted with bovine RSV; G gene sequences unrooted. Sequence symbols are colour coded by geography to emphasize the congruence between the phylogenies of the two genome regions. Monophyletic groupings which contain sequences from the 07/08 respiratory season for which sequence data was available from all four referral centres are indicated in shaded boxes. Bootstrap values >70% are indicated.

doi:10.1371/journal.pone.0017427.g002

UK and Belgium; these may differentially evolve and/or be under dissimilar structural or immunologic constraints. Alternatively, we may have identified more positively selected sites due to the larger dataset analysed. The HRSV residues identified as under positive selection 142, 206, 274 and 286 have been associated with substitutions in successful antibody escape mutants [68–70]. The positively selected residues 215, 217 and 226 fall within a region thought to be immunogenic of neutralising antibodies [71,72], and residue 297 has previously been identified as a determinant of the integrity of multiple overlapping strain-specific epitopes [73].

An inability of antibodies to select for mutations at sites autonomous to their binding specificity has previously been used to support the notion that HRSV-G epitopes are linear rather than conformational [45]. The greatest distance between any two residues identified as under positive selection within HRSV-A, excluding the conserved region between residues 151–190 (in which we identified one residue under positive selection at position 161), was eight amino acids, which suggests that HRSV-A might potentially generate variability in any epitope of the G protein. The four residues identified as under positive selection in the previous study which were not verified through this work were nevertheless proximal to residues identified as positively selected in this study, with the greatest distance between the two being four amino acids. This supports the notion that within epitopes, changes in any residue might be selected for immune evasion.

For HRSV-B, only five residues were identified as under positive selection, compared with twelve during a previous study. Again, we analysed a greater number of sequences than previous work, though both studies used sequences generated from samples collected over a similar time frame [11]. Here, all the identified sites under positive selection were in the second hyper-variable region in the ectodomain, and only two of the five residues we identified as positively selected were also identified as positively selected previously. Two of the discordant sites we identified were within or downstream of the previously identified [11,43] 60 nucleotide repeat insertion at the 3′-proximal end of the G protein gene, and so it is possible that these sites were not detected by previous analyses due to limited availability of sequence data for the insert region. Residue 224, the third amino acid not previously identified as such, was determined as positively selected with a probability of $p = 0.501$, and was not detected by the Naïve Empirical Bayes test (used in previous analyses), accounting for the discordancy at this residue between studies.

Six-nucleotide in frame deletions at amino acid positions 159 and 160 reported previously [11] were observed in four of the Newcastle HRSV isolates from 1985, 1986 and 2009. This occurred in different bootstrap-supported lineages of the G region phylogenetic tree, providing strong evidence that this deletion has been independently selected for more than once. Recent identification of two epitopes within the central conserved region of the HRSV G protein ectodomain between residues 151–163 and 164–176 [49] illustrates the immunogenicity of these two peptide regions, while an investigation of the properties of the central conserved domain of HRSV-G showed that the region between amino acids 149–177 played no role in

virus infectivity [74]. A loss of these two residues may therefore reduce virus immunogenicity while having no effect on virus infectivity. These observations, together with previous reports of premature stop codons and frame shifts within the subgroup B G protein [71,73,75], suggest that HRSV-B may use quite different mechanisms from HRSV-A to evade host immune responses.

A number of residues were identified as positively selected within the G protein of HMPV types A, B1 and B2 (14, 6 and 17) which were differentially located between lineages, in keeping with the observations made of HRSV. There is a predicted cytotoxic T cell epitope between residues 32–41 [76], and within this region one site was identified as positively selected within HMPV-A. The residue identified as under positive selection in all three HMPV-G analyses (residue 105) is not a predicted site of N- or O-glycosylation [51].

Phylogenetic analyses of HRSV and HMPV yielded the common observation that strains isolated from geographically widespread referral centres frequently resolved within the same lineage, reflecting the ability of these viruses to disseminate rapidly on a global scale, and substantiating previous reports of the worldwide distribution of lineages of both viruses [11,63,68,77].

Switching of the predominantly circulating genotype (HMPV)/subtype (HRSV) in a population is widely discussed, but poorly understood in terms of what drives these events or the mechanisms by which they occur. The MRCA estimates for contemporary circulating clades were for HRSV and HMPV 14–17 and 11–28 years respectively, providing conclusive evidence that switching of the predominantly circulating genetic group of both viruses arises independently of novel lineage emergence events. The differences in evolutionary rates between older and more recent HMPV isolates, interpreted here as evidence of an increasing population size, contradicts a previous analysis which showed that one HMPV lineage was increasing in size whereas another was decreasing [57]. Taken together, this information lends to the hypothesis that HMPV (and HRSV) lineages circulate in a cyclic trend of multiple eclipse phases preceding periodic population expansions. Proliferation occurs when the lineage is of minimal susceptibility to the adaptive immune responses of the host population, and a regression in circulating frequency occurs as the host population is increasingly exposed. During the eclipse phase, the virus evolves immune evasive characteristics, which when accumulated sufficiently permit a new phase of widespread circulation.

In summary, we have analysed the molecular epidemiology and evolution of HRSV and HMPV in parallel using the novel approach of clade identification for evolutionary analysis. This work has revealed a number of shared trends, including evidence of both locally and globally circulating lineages of both viruses, significant positive selection acting in the G but not the F genes and a lack of evidence for positive selection being restricted to specific codons. Switching of the predominantly circulating subtype (HRSV) or genotype (HMPV) may be a result of fluctuating circulating frequencies of contemporary clades, which cycle through proliferative and eclipse phases, and is not due to novel lineage emergence events. We suggest that HRSV has the ability to select for residue substitutions at multiple sites within

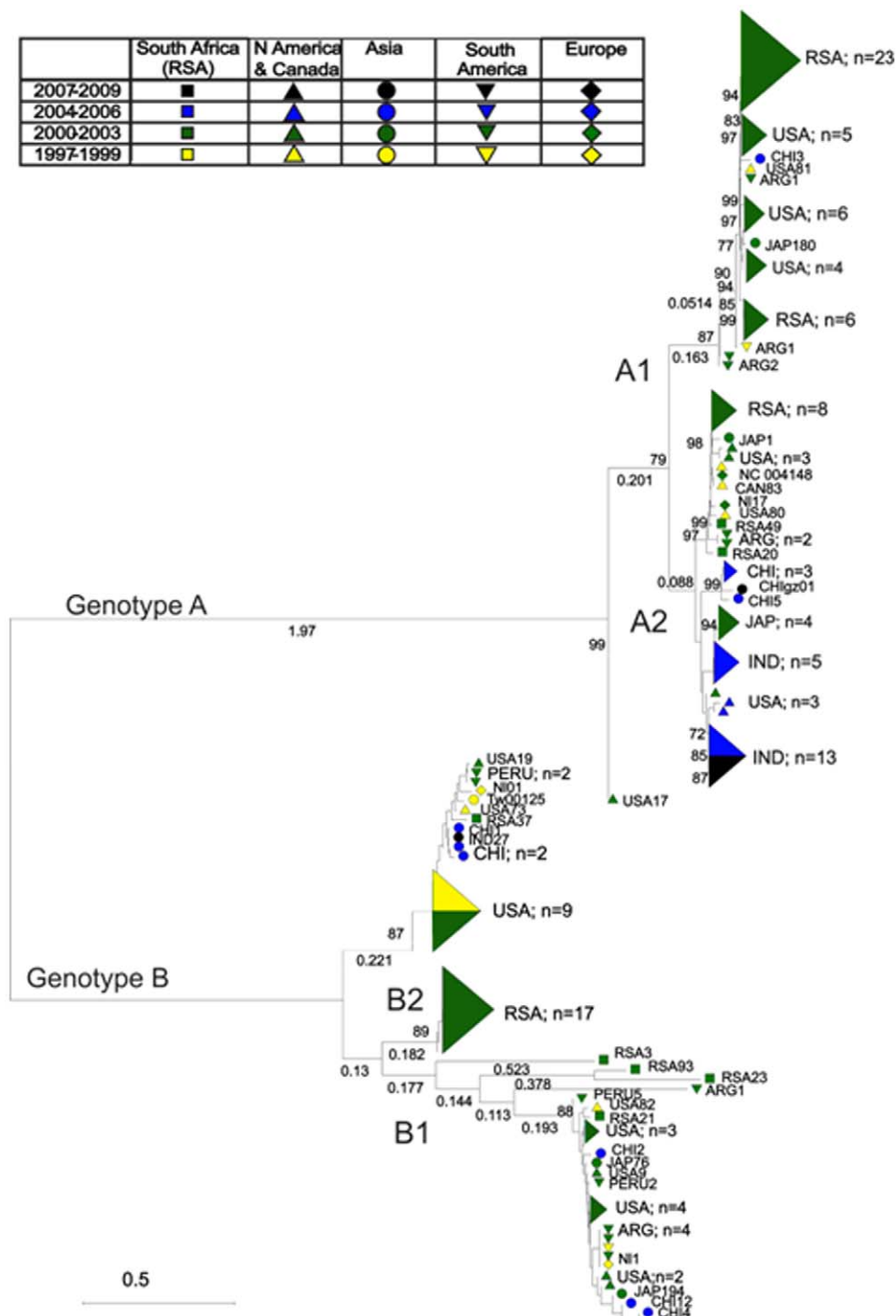


Figure 3. Phylogenetic analysis of HMPV partial G gene sequences (unrooted). Phylogenetic reconstruction was by neighbour joining of MCL-corrected pair-wise distances. Genotypes (A and B) and sub-genotypes (A1, A2, B1 and B2) are indicated. Sequence symbols are colour coded by year of isolation and symbol shape is designated depending on geographic origin of sequence. Bootstrap values >70% are indicated. doi:10.1371/journal.pone.0017427.g003

epitopes, contributing to the successful recirculation to high incidence of lineages of this virus.

Methods

Sample collection

HRSV and HMPV positive respiratory samples archived between March 2006 and December 2008 at the Specialist

Virology Centre (SVC), Royal Infirmary of Edinburgh, UK were identified as described previously [13]. 26 HRSV isolates were collected between 1965 and 2009 from Newcastle, UK [6]. Additionally, nine HRSV and eight HMPV positive samples from Bangkok, Thailand (2006–07 respiratory season) detected as described previously were included in analyses [78] along with 16 HRSV and 16 HMPV variants from the Academic Medical Centre, Amsterdam, Netherlands from the 2007–8 season.

HRSV and HMPV nucleotide amplification and sequencing in the F and G genes

All HRSV ($n=183$) and HMPV ($n=177$) positive respiratory samples were amplified and sequenced in the 3' F gene region as previously described [13] (589 and 438 nucleotides respectively). For HRSV, all 26 isolates from Newcastle, seven from Bangkok, 16 from the Netherlands and eight from the UK were amplified in the 3' G gene region that included both HVRs (780 nucleotides). Combined, the HRSV variants analysed were globally distributed and spanned 44 years. These were analysed alongside available HMPV G gene sequences in Genbank, which encompassed a temporal diversity of 12 years.

HRSV and HMPV RNA was extracted using Qiagen QIAamp viral RNA mini kit and reverse transcribed using Qiagen A3500 reverse transcription system, with extended elongation of 55 minutes and use of random primers. HRSV and HMPV cDNAs were amplified by nested PCR. Reaction mixtures contained 4 μ l MgCl₂ (25 mM), 0.2 μ l dNTP (3 mM), 1 μ l each outer primer (10 mM) and 0.08 μ l TaqPolymerase. Primers for HRSV-F gene PCR were (outer sense) 916- TAT GGW GTD ATA GAY ACM CCY TGY TGG, (inner sense) 1018- GG RTG GTA YTG TGA YAA TGC AGG, (inner antisense) 1663-CT TAR TGT RAC TGG TGT GYT TYT GGC and (outer antisense) 1682- TWC CAC TYA GTT GRT CYT TRC TTA RTG. HRSV-G gene was amplified by primers (outer sense) 47- CCT GGG AYA CTC TYA ATC AT, (inner sense) 137-TGG CAA TGA TAA TCT CAA C (inner antisense) 117- CCT YTG CTA ACT GCA CT and (outer antisense) 147-GTA TAC CAA CCW GTT CTT A; antisense primers align in the downstream fusion gene. Primers for HMPV F gene amplification were as described previously [13]. 2 μ l cDNA was used in the first round and 1 μ l first round product was used in the second round reaction. The same cycling conditions were used throughout; 30 cycles at 94°C for 18 s, 50°C for 90 s and 72°C for 30 s, and a terminal 72°C elongation step for 300 s.

Sequences obtained in the course of this study have been submitted to GenBank and assigned accession numbers GU386461–GU386756 (HMPV and HRSV F gene) and HQ731687–HQ731784 (HRSV G gene).

Phylogenetic analysis

A summary of the computational techniques undertaken for this work and the sequence datasets analysed is tabulated (Table 1). Partial F gene sequences were aligned and genetic distances were calculated using Simmonics v1.9 sequence editor package (www.virus-evolution.org). Phylogenetic trees were constructed from 1000 samplings of maximum composite likelihood (MCL) distances by neighbour-joining method with pair-wise deletions for missing nucleotides in MEGA v4.0. For HRSV, 58 isolates were available for sequencing in both F and G gene regions, and these subsets were phylogenetically analysed separately for comparison using the same methods. HMPV-G sequences downloaded from Genbank were analysed altogether (unrooted). The dataset parameters used for phylogenetic reconstructions are summarized (Table 2).

Identification of HRSV-F and HMPV-F clades

No systematic method is currently used for identification of distinct HRSV and HMPV lineages. HRSV and HMPV clades (defined as described herein) were identified for the purpose of evolutionary modelling. Phylogenetic analyses of F gene nucleotide sequences of contemporary HRSV and HMPV strains (generated from samples collected since 2007) were constructed as described above for identification of bootstrap-supported monophyletic lineages (values $\geq 70\%$) using MEGA v4.0. Subsequent phylogenetic analyses incorporated older monophyletic sequences within the contemporary clades, with visually appropriate limitations of 1.5–3% variation across clades at the nucleotide level and no individual sequence varying from all others within the clade by more than 0.5% at the nucleotide level.

Table 1. Summary of the evolutionary analyses undertaken by taxonomic group.

Taxonomic group	Gene	Analyses undertaken (software used)			
		Nucleotide sequencing (Simmonics)	Phylogenetic (MEGA)	Positive selection (PAML)	Evolutionary (BEAST)
HRSV <i>spp</i>	F	×	×		
HRSV <i>spp</i>	G	×	×		
HRSV-A	F	×		×	
HRSV-B	F	×		×	
HRSV-A	G	×		×	
HRSV-B	G	×		×	
HRSV clades	F	×			×
HMPV <i>spp</i>	F	×	×		
HMPV <i>spp</i>	G		×		
HMPV-A2	F	×		×	
HMPV-B1	F	×		×	
HMPV-B2	F	×		×	
HMPV-A	G			×	
HMPV-B1	G			×	
HMPV-B2	G			×	
HMPV clades	F	×			×

doi:10.1371/journal.pone.0017427.t001

Table 2. Sequence datasets for phylogenetic and for positive selection analyses.

Virus group	Gene	Region analysed	Reference strain	Root for phylogenetic reconstruction
HRSV-A	G	Residue 13 to C terminus	Long	HRSV-B
HRSV-B	G	Residue 35 to C terminus	WV/B1/85	HRSV-A
HRSV	F	Residues 358–554	Long	Bovine RSV
HMPV-A1	G	Residue 1 to C terminus	Analysed for positive selection only	
HMPV-B	G	Residue 35 to C terminus	Analysed for positive selection only	
HMPV	G	Residue 35 to C terminus	CAN97-83	Unrooted
HMPV	F	Residues 385–531	CAN97-83	AMPV-C

doi:10.1371/journal.pone.0017427.t002

Identification of positively selected sites in the F protein of HMPV and HRSV clades and the G protein of genotypes/subtypes

Prior to evolutionary analyses, positively selected sites were removed from nucleotide alignments. This was considered necessary as positively selected sites undergo convergent evolution whereas other sites are subject to neutral or nearly neutral drift, and different evolutionary mechanisms inevitably violate assumptions of the SRD06 evolutionary model [59]. Nucleotide alignments of HRSV and HMPV subtypes/sub-genotypes were analysed for positive selection using PAML v4.4. In the attachment protein of HRSV, species level analyses detected all sites in the ectodomain as positively selected due to the high diversity in this region, and similar results were produced for the HMPV species level analyses. This was likely also affected by the large amount of sequence data available. To conserve the maximum diversity in the sequence dataset analysed, analyses for positive selection were undertaken in decreasing increments of taxonomic diversity until satisfactory results were attained. This was found to be the subtype level for HRSV and the sub-genotype level for HMPV. The recommended combination of models 0, 1a and 2a [79] were run and Bayes empirical Bayes results only were considered as recommended. The genome regions analysed for the different virus subgroups are summarized (Table 2). Nucleotide alignments with positively selected sites

removed were reanalysed in PAML to confirm no positive selection was detected.

Clade turnover

Identical sequences (by geography, date and nucleotide sequence) within clades were removed, and evolutionary analyses were restricted to clades with a minimum of 15 non-identical sequences. Three HRSV and five HMPV clades were identified (indicated on Fig. 1). Evolutionary analyses of HRSV and HMPV clades were undertaken using BEAST to calculate evolutionary rates and time since the most recent common ancestor (*tMRC*) of genetic groups. The F gene region was selected as it was more phylogenetically informative, not subject to positive selection and sequences from a wider geographical and temporal range were available. Clade sequence datasets were run in BEAST using a strict SRD06 model, which allows the third position in a codon to have a different substitution rate to the first and second, until all ESS (expected sample size) values exceeded 200 (recommended). The strict model, as opposed to a relaxed model, assumes that all lineages incorporated within the sequence dataset evolve at the same rate. As the sequence datasets described here were by definition monophyletic groups, the assumption of a nonvariable evolutionary rate within each group was justified. Analyses were run in duplicate to ensure convergence of the posterior distribution, demonstrating repeatability of the result. The coefficient of variation histogram was used to confirm validity of the strict model.

Table 3. Positively selected sites detected in the attachment (G) protein of HRSV and HMPV.

Group	Reference strain	Positively selected sites	Frequency of positively selected sites identified in probability range				
			0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–1.0
HRSV-A	Long	101, 104, 106, 111, 115, 117, 121, 122, 123, 126, 127, 131, 142, 146, 161, 206, 215, 217, 226, 230, 233, 250, 258, 262, 274, 276, 280, 286, 289, 290, 291, 297	13	4	5	5	5
HRSV-B	WV/B1/85	223, 224, 258, 267/287 ¹ , 297/317	1	2		1	1
HMPV-A	CAN97-83	33, 81, 84, 93, 105, 106, 110, 145, 146, 157, 165, 172, 177, 190	3	2	3	5	1
HMPV-B1	CAN97-83	70, 100, 105, 116, 162, 201		2	3		1
HMPV-B2	CAN97-83	53, 55, 85, 89, 93, 105, 109, 111, 121, 126, 137, 141, 180, 202, 207, 217, 222	4	8	3	1	1

¹Refers to strains without/with the 60 nt repeated region.

doi:10.1371/journal.pone.0017427.t003

Table 4. Substitution rates and estimates of t_{MRC} for HRSV and HMPV subtypes and clades, calculated using F gene sequences.

Virus	Range of isolation dates (yrs)	Genetic dist (%)	dN/dS	Evolutionary rate $\times 10^{-4}$ (95% HPD)	t_{MRC} (95% HPD)
HRSV clade 1	2	1.53	0.159	ND ¹	ND ¹
HRSV clade 2	15	2.38	0.0252	13.2 (5.84–20.9)	17 (15–22)
HRSV clade 3	12	2.89	0.0746	15.5 (8.89–22.2)	14 (12–17)
HMPV clade 1	24	2.51	0.104	10.4 (5.50–15.6)	28 (24–33)
HMPV clade 2	11	2.74	0.0746	17.3 (7.14–28.2)	13 (11–16)
HMPV clade 3	6	2.05	0.0212	13.1 (1.19–28.9)	11 (6–22)
HMPV clade 4	26	2.06	0.0705	12.9 (5.75–20.8)	27 (25–30)
HMPV clade 5	16	2.97	0.0229	11.1 (6.38–16.0)	19 (16–24)

¹Not Determined.

doi:10.1371/journal.pone.0017427.t004

Acknowledgments

We are very grateful to Peter McCulloch, Mary Notman, Julie White and Carol Thomson (Specialist Virology Centre, Royal Infirmary of Edinburgh, UK) for providing samples and virus testing results from the respiratory sample archive in Edinburgh, and to Professor Paul Sharp (University of Edinburgh, UK) for helpful discussions when preparing the manuscript.

Ethical approval. Lothian Regional Ethics Committee (08/S11/02/2).

References

- Anderson LJ, Hierholzer JC, Tsou C, Hendry RM, Fernie BF, et al. (1985) Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J Inf Dis* 151(4): 626–633.
- Mufson MA, Orvell C, Rafnar B, Norrby E (1985) Two distinct subtypes of human respiratory syncytial virus. *J Gen Virol* 66(10): 2111–2124.
- van den Hoogen BG, de Jong JC, Groen J, Kuiken T, de Groot R, et al. (2001) A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nat Med* 7: 719–724.
- McGill A, Greensill J, Marsh R, Craft AW, Toms GL (2004) Detection of human respiratory syncytial virus genotype specific antibody responses in infants. *J Med Virol* 74(3): 492–498.
- Venter M, Madhi SA, Tiemessen CT, Schoub BD (2001) Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. *J Gen Virol* 82(9): 2117–2124.
- Morgan LA, Routledge EG, Willcocks MM, Samson ACR, Scott R, et al. (1987) Strain variation of respiratory syncytial virus. *J Gen Virol* 68(11): 2781–2788.
- Taylor CE, Morrow S, Scott M, Young B, Toms GL (1989) Comparative virulence of respiratory syncytial virus subgroups A and B. *Lancet* 333(8641): 777–778.
- Cane PA, Matthews DA, Pringle CR (1994) Analysis of respiratory syncytial virus strain variation in successive epidemics in one city. *J Clin Microbiol* 32(1): 1–4.
- Peret T, Hall C, Hammond G, Piedra P, Storch G, et al. (2000) Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. *J Inf Dis* 181(6): 1891–1896.
- Zlateva KT, Lemey P, Vandamme AM, Van Ranst M (2004) Molecular evolution and circulation patterns of human respiratory syncytial virus subgroup A: positively selected sites in the attachment G glycoprotein. *J Virol* 78(9): 4675–4683.
- Zlateva KT, Lemey P, Moes E, Vandamme AM, Van Ranst M (2005) Genetic variability and molecular evolution of the human respiratory syncytial virus subgroup B attachment G protein. *J Virol* 79(14): 9157–9167.
- Agapov E, Sumino KAC, Gaudreault-Keener M, Storch GAA, Holtzman MAJ (2006) Genetic variability of human metapneumovirus infection: evidence of a shift in viral genotype without a change in illness. *J Inf Dis* 193(3): 396–403.
- Gaunt E, McWilliam-Leitch EC, Templeton K, Simmonds P (2009) Incidence, molecular epidemiology and clinical presentations of human metapneumovirus; assessment of its importance as a diagnostic screening target. *J Clin Virol* 46(4): 318–324.
- Åkerlind B, Norrby E (1986) Occurrence of respiratory syncytial virus subtypes a and b strains in sweden. *J Med Virol* 19(3): 241–247.
- Deng J, Qian Y, Zhu RN, Wang F, Zhao LQ (2006) Surveillance for respiratory syncytial virus subtypes A and B in children with acute respiratory infections in Beijing during 2000 to 2006 seasons [Article in Chinese]. *Zhonghua Er Ke Za Zhi* 44(12): 924–7.
- Hendry RM, Pierik LT, McIntosh K (1989) Prevalence of respiratory syncytial virus subgroups over six consecutive outbreaks: 1981–1987. *J Inf Dis* 160(2): 185–190.
- Mufson MA, Belshe RB, Orvell C, Norrby E (1988) Respiratory syncytial virus epidemics: variable dominance of subgroups A and B strains among children, 1981–1986. *J Inf Dis* 157(1): 143–8.
- Freymuth F, Petitjean J, Pothier P, Brouard J, Norrby E (1991) Prevalence of respiratory syncytial virus subgroups A and B in France from 1982 to 1990. *J Clin Microbiol* 29(3): 653–655.
- Monto AS, Ohmit S (1990) Respiratory syncytial virus in a community population: circulation of subgroups A and B since 1965. *J Inf Dis* 161(4): 781–783.
- Zlateva KT, Vijgen L, Dekeersmaecker N, Naranjo C, Van Ranst M (2007) Subgroup prevalence and genotype circulation patterns of human respiratory syncytial virus in Belgium during 10 successive epidemic seasons. *J Clin Microbiol*. JCM.00339-07.
- Matsuzaki Y, Itagaki T, Abiko C, Aoki Y, Suto A, et al. (2008) Clinical impact of human metapneumovirus genotypes and genotype-specific seroprevalence in Yamagata, Japan. *J Med Virol* 80(6): 1084–9.
- Vicente D, Montes M, Cilla G, Perez-Yarza EG, Perez-Trallero E (2006) Differences in clinical severity between genotype A and genotype B human metapneumovirus infection in children. *Clin Inf Dis* 42(12): e111–3.
- Bastien N, Ward D, Van Caeseele P, Brandt K, Lee SHS, et al. (2003) Human metapneumovirus infection in the Canadian population. *J Clin Microbiol* 41(10): 4642–4646.
- Boivin G, Abed Y, Pelletier G, Ruel L, Moisan D, et al. (2002) Virological features and clinical manifestations associated with human metapneumovirus: a new paramyxovirus responsible for acute respiratory tract infections in all age groups. *J Inf Dis* 186(9): 1330–1334.
- Esper F, Boucher D, Weibel C, Martinello RA, Kahn JS (2003) Human metapneumovirus infection in the United States: clinical manifestations associated with a newly emerging respiratory infection in children. *Pediatrics* 111(6): 1407–1410.
- Falsey AR, Erdman D, Anderson LJ, Walsh EE (2003) Human metapneumovirus infections in young and elderly adults. *J Inf Dis* 187(5): 785–790.
- Stockton J, Stephenson I, Fleming D, Zambon M (2002) Human metapneumovirus as a cause of community-acquired respiratory illness. *Emerg Inf Dis* 8(9): 897–901.
- Hendry RM, Burns JC, Walsh EE, Graham BS, Wright PF, et al. (1988) Strain-specific serum antibody responses in infants undergoing primary infection with respiratory syncytial virus. *J Inf Dis* 157(4): 640–647.
- Olmsted RA, Elango N, Prince GA, Murphy BR, Johnson PR, et al. (1986) Expression of the F glycoprotein of respiratory syncytial virus by a recombinant vaccinia virus: comparison of the individual contributions of the F and G glycoproteins to host immunity. *Proc Nat Acad Sci USA* 83(19): 7462–7466.

30. Endo R, Ebihara T, Ishiguro N, Teramoto S, Ariga T, et al. (2008) Detection of four genetic subgroup-specific antibodies to human metapneumovirus attachment (G) protein in human serum. *J Gen Virol* 89(8): 1970–1977.
31. Skiadopoulos MH, Biacchesi S, Buchholz UJ, Riggs JM, Surman SR, et al. (2004) The two major human metapneumovirus genetic lineages are highly related antigenically, and the fusion (F) protein is a major contributor to this antigenic relatedness. *J Virol* 78(13): 6927–6937.
32. van den Hoogen BG, Herfst S, de Graaf M, Sprong L, van Lavieren R, et al. (2007) Experimental infection of macaques with human metapneumovirus induces transient protective immunity. *J Gen Virol* 88(4): 1251–1259.
33. Herfst S, de Graaf M, Schrauwen EJA, Ulbrandt ND, Barnes AS, et al. (2007) Immunization of Syrian golden hamsters with F subunit vaccine of human metapneumovirus induces protection against challenge with homologous or heterologous strains. *J Gen Virol* 88(10): 2702–2709.
34. Hernandez LD, Hoffman LR, Wolfsberg TG, White JM (1996) Virus-cell and cell-cell fusion 1. *Ann Rev Cell and Dev Biol* 12(1): 627–661.
35. Chan DC, Kim PS (1998) HIV entry and its inhibition. *Cell* 93(5): 681–684.
36. Johnson PR, Olmsted RA, Prince GA, Murphy BR, Alling DW, et al. (1987) Antigenic relatedness between glycoproteins of human respiratory syncytial virus subgroups A and B: evaluation of the contributions of F and G glycoproteins to immunity. *J Virol* 61(10): 3163–3166.
37. Cane PA (1997) Analysis of linear epitopes recognised by the primary human antibody response to a variable region of the attachment (G) protein of respiratory syncytial virus. *J Med Virol* 51(4): 297–304.
38. Palomo C, Cane PA, Melero JA (2000) Evaluation of the antibody specificities of human convalescent-phase sera against the attachment (G) protein of human respiratory syncytial virus: Influence of strain variation and carbohydrate side chains. *J Med Virol* 60(4): 468–474.
39. Johnson PR, Spriggs MK, Olmsted RA, Collins PL (1987) The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proc Nat Acad Sci USA* 84(16): 5625–5629.
40. Ishiguro N, Ebihara T, Endo R, Ma X, Kikuta H, et al. (2004) High genetic diversity of the attachment (G) protein of human metapneumovirus. *J Clin Microbiol* 42(8): 3406–3414.
41. MacArthur WM, Thornton JM (1991) Influence of proline residues on protein conformation. *Journal of Molecular Biology* 218(2): 397–412.
42. Wertz GW, Collins PL, Huang Y, Gruber C, Levine S, et al. (1985) Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein. *Proc Nat Acad Sci USA* 82(12): 4075–4079.
43. Botosso VF, Zanotto PMde A, Ueda M, Arruda E, Gilio AE, et al. (2009) Positive selection results in frequent reversible amino acid replacements in the G protein gene of human respiratory syncytial virus. *PLoS Pathog* 5(1): e1000254.
44. Langedijk JPM, Schaaper WMM, Meloen RH, van Oirschot JT (1996) Proposed three-dimensional model for the attachment protein G of respiratory syncytial virus. *J Gen Virol* 77(6): 1249–1257.
45. Melero JA, Garcia-Barreno B, Martinez I, Pringle CR, Cane PA (1997) Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. *J Gen Virol* 78(10): 2411–2418.
46. Feldman SA, Hendry RM, Beeler JA (1999) Identification of a linear heparin binding domain for human respiratory syncytial virus attachment glycoprotein G. *J Virol* 73(8): 6610–6617.
47. Lerch RA, Anderson K, Wertz GW (1990) Nucleotide sequence analysis and expression from recombinant vectors demonstrate that the attachment protein G of bovine respiratory syncytial virus is distinct from that of human respiratory syncytial virus. *J Virol* 64(11): 5559–5569.
48. Tripp RA, Jones LP, Haynes LM, Zheng H, Murphy PM, et al. (2001) CX3C chemokine mimicry by respiratory syncytial virus G glycoprotein. *Nat Immunol* 2(8): 732–8.
49. Murata Y, Lightfoote PM, Falsey AR, Walsh EE (2010) Identification and human serum reactivity of neutralizing epitopes within the central unglycosylated region of the respiratory syncytial virus attachment protein. *Clin Vacc Immunol* 17(4): 695–7.
50. Biacchesi S, Skiadopoulos MH, Boivin G, Hanson CT, Murphy BR, et al. (2003) Genetic diversity between human metapneumovirus subgroups. *Virology* 315(1): 1–9.
51. Peret TCT, Abed Y, Anderson LJ, Erdman DD, Boivin G (2004) Sequence polymorphism of the predicted human metapneumovirus G glycoprotein. *J Gen Virol* 85(3): 679–686.
52. Bukreyev A, Yang L, Fricke J, Cheng L, Ward JM, et al. (2008) The secreted form of the G glycoprotein of respiratory syncytial virus helps the virus evade antibody-mediated restriction of replication by acting as an antigen decoy and through effects on Fc receptor-bearing leukocytes. *J Virol* 82(10): 4604–4608.
53. Scott PD, Ochola R, Ngama M, Okiro EA, Nokes D, et al. (2006) Molecular analysis of respiratory syncytial virus reinfections in infants from coastal Kenya. *J Inf Dis* 193(1): 59–67.
54. White LJ, Waris M, Cane PA, Nokes DJ, Medley GF (2005) The transmission dynamics of groups A and B human respiratory syncytial virus (hRSV) in England & Wales and Finland: seasonality and cross-protection. *Epidemiol Infect* 133(2): 279–89.
55. Reitter JN, Means RE, Desrosiers RC (1998) A role for carbohydrates in immune evasion in AIDS. *Nat Med* 4(6): 679–84.
56. de Graaf M, Osterhaus ADME, Fouchier RAM, Holmes EC (2008) Evolutionary dynamics of human and avian metapneumoviruses. *J Gen Virol* 89(12): 2933–2942.
57. Padhi A, Verghese B (2008) Positive natural selection in the evolution of human metapneumovirus attachment glycoprotein. *Virus Res* 131(2): 121–131.
58. Yang CF, Wang C, Tollefson S, Piyaratna R, Lintao L, et al. (2009) Genetic diversity and evolution of human metapneumovirus fusion protein over twenty years. *J Virol* 83(1): 138.
59. Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7(1): 214.
60. McWilliam Leitch EC, Cabrero M, Cardosa J, Harvala H, Ivanova OE, et al. (2010) Evolutionary dynamics and temporal/geographical correlates of recombination in the human enteroviruses, echovirus 9, 11 and 30. *J Virol* 84(10): 5673–5680.
61. Zhang C, Mammen MP, Jr., Chinnawirotpisan P, Klungthong C, Rodpradit P, et al. (2005) Clade replacements in dengue virus serotypes 1 and 3 are associated with changing serotype prevalence. *J Virol* 79(24): 15123–15130.
62. Santibanez S, Tischer A, Heider A, Siedler A, Hengel H (2002) Rapid replacement of endemic measles virus genotypes. *J Gen Virol* 83(11): 2699–2708.
63. Garcia O, Martin M, Dopazo J, Arbiza J, Frabasile S, et al. (1994) Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. *J Virol* 68(9): 5448–5459.
64. Sullender WM, Mufson MA, Anderson LJ, Wertz GW (1991) Genetic diversity of the attachment protein of subgroup B respiratory syncytial viruses. *J Virol* 65(10): 5425–5434.
65. Sullender WM, Mufson MA, Prince GA, Anderson LJ, Wertz GW (1998) Antigenic and genetic diversity among the attachment proteins of group A respiratory syncytial viruses that have caused repeat infections in children. *J Inf Dis* 178(4): 925–932.
66. Woelk CH, Holmes EC (2001) Variable immune-driven natural selection in the attachment (G) glycoprotein of respiratory syncytial virus (RSV). *J Mol Evol* 52(2): 182–192.
67. Lopez JA, Andreu D, Carreno C, Whyte P, Taylor G, et al. (1993) Conformational constraints of conserved neutralizing epitopes from a major antigenic area of human respiratory syncytial virus fusion glycoprotein. *J Gen Virol* 74(12): 2567–2577.
68. Cane PA, Pringle CR (1995) Evolution of subgroup A respiratory syncytial virus: evidence for progressive accumulation of amino acid changes in the attachment protein. *J Virol* 69(5): 2918–2925.
69. Martinez I, Dopazo J, Melero JA (1997) Antigenic structure of the human respiratory syncytial virus G glycoprotein and relevance of hypermutation events for the generation of antigenic variants. *J Gen Virol* 78(10): 2419–2429.
70. Walsh EE, Falsey AR, Sullender WM (1998) Monoclonal antibody neutralization escape mutants of respiratory syncytial virus with unique alterations in the attachment (G) protein. *J Gen Virol* 79(3): 479–487.
71. Garcia-Barreno B, Portela A, Delgado T, Lopez JA, Melero JA (1990) Frame shift mutations as a novel mechanism for the generation of neutralization resistant mutants of human respiratory syncytial virus. *EMBO J* 9: 4181–4187.
72. Olmsted RA, Murphy BR, Lawrence LA, Elango N, Moss B, et al. (1989) Processing, surface expression, and immunogenicity of carboxy-terminally truncated mutants of G protein of human respiratory syncytial virus. *J Virol* 63(1): 411–420.
73. Rueda P, Palomo CN, Garcia-Barreno B, Melero JA (1995) The three C-terminal residues of human respiratory syncytial virus G glycoprotein (Long strain) are essential for integrity of multiple epitopes distinguishable by antidiotype antibodies. *Viral Immunol* 8(1): 37–46.
74. Gorman JJ, McKimm-Breschkin JL, Norton RS, Barnham KJ (2001) Antiviral activity and structural characteristics of the nonglycosylated central subdomain of human respiratory syncytial virus attachment (G) glycoprotein. *J Biol Chem* 276(42): 38988–38994.
75. Martinez I, Valdes O, Delfraro A, Arbiza J, Russi J, et al. (1999) Evolutionary pattern of the G glycoprotein of human respiratory syncytial viruses from antigenic group B: the use of alternative termination codons and lineage diversification. *J Gen Virol* 80(1): 125–130.
76. Herd KA, Mahalingam S, Mackay IM, Nissen M, Sloots TP, et al. (2006) Cytotoxic T-lymphocyte epitope vaccination protects against human metapneumovirus infection and disease in mice. *J Virol* 80(4): 2034–2044.
77. Boivin G, Mackay I, Sloots TP, Madhi S, Freymuth F, et al. (2004) Global genetic diversity of human metapneumovirus fusion gene. *Emerg Inf Dis* 10(6): 1154–7.
78. Chieochansin T, Samransamruajkit R, Chutinimitkul S, Payungporn S, Hiranras T, et al. (2008) Human bocavirus (HBoV) in Thailand: clinical manifestations in a hospitalized pediatric patient and molecular virus characterization. *J Infect* 56(2): 137–142.
79. Yang Z, Wong WSW, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol* 22(4): 1107–1118.